

SEMIDEFINITE PROGRAMMING RELAXATIONS FOR LINEAR SEMI-INFINITE POLYNOMIAL PROGRAMMING

FENG GUO

ABSTRACT. This paper studies a class of so-called linear semi-infinite polynomial programming (LSIPP) problems. It is a subclass of linear semi-infinite programming problems whose constraint functions are polynomials in parameters and index sets are basic semialgebraic sets. When the index set of an LSIPP problem is compact, a convergent hierarchy of semidefinite programming (SDP) relaxations is constructed under the assumption that the Slater condition and the Archimedean property hold. When the index set is noncompact, we use the technique of homogenization to equivalently convert the LSIPP problem into compact case under some generic assumption. Consequently, a corresponding hierarchy of SDP relaxations for noncompact LSIPP problems is obtained. We apply this relaxation approach to the special LSIPP problem reformulated from a polynomial optimization problem. A new SDP relaxation method is derived for solving the class of polynomial optimization problems whose objective polynomials are stably bounded from below on noncompact feasible sets.

Key words linear semi-infinite programming, semidefinite programming relaxations, sum of squares, polynomial optimization

1. INTRODUCTION

We consider the following *linear semi-infinite polynomial programming* (LSIPP) problem

$$(1.1) \quad (P) \quad \begin{cases} p^* := \inf_{x \in \mathbb{R}^m} c^T x \\ \text{s.t. } a(y)^T x + b(y) \geq 0, \quad \forall y \in S \subseteq \mathbb{R}^n, \end{cases}$$

where $c \in \mathbb{R}^m$, $b(Y) \in \mathbb{R}[Y] := \mathbb{R}[Y_1, \dots, Y_n]$ the polynomial ring in Y over the real field, $a(Y) = (a_1(Y), \dots, a_m(Y))^T \in \mathbb{R}[Y]^m$, and the *index set* S is a basic semialgebraic set defined by

$$(1.2) \quad S := \{y \in \mathbb{R}^n \mid g_1(y) \geq 0, \dots, g_s(y) \geq 0\},$$

where $g_j(Y) \in \mathbb{R}[Y]$, $j = 1, \dots, s$. In this paper, we assume that (1.1) is feasible and bounded from below, i.e., $-\infty < p^* < \infty$. Note that the problem (1.1) is NP-hard. Indeed, it is obvious that the problem of minimizing a polynomial $f(Y) \in \mathbb{R}[Y]$ over S can be regarded as a special LSIPP problem (see Section 4). As is well known, the polynomial optimization problem is NP-hard even when $n > 1$, $f(Y)$ is a nonconvex quadratic polynomial and $g_j(Y)$'s are linear [23]. Hence, a general LSIPP problem can not be expected to be solved in polynomial time unless P=NP.

LSIPP can be seen as a special branch of *linear semi-infinite programming* (LSIP), or more general, of *semi-infinite programming* (SIP), in which the involved

functions are not necessarily polynomials. Due to its many applications and appealing theoretical properties, SIP has become an independent and active research area since 1960s and a large amount of work has been done on it. Numerically, SIP problems can be solved by different approaches including, for instance, discretization method, local reduction method, exchange method, simplex-like method and so on. See the surveys [5, 6, 7, 11, 16, 27, 30] and the references therein for details. To the best of our knowledge, few of them are specially designed by exploiting features of polynomial optimization problems. Parpas and Rustem [24] proposed a discretization-like method to solve min-max polynomial optimization problems, which can be reformulated as *semi-infinite polynomial programming* (SIPP) problems. Using polynomial approximation and an appropriate hierarchy of *semidefinite programming* (SDP) relaxations, Lasserre presented an algorithm to solve the *generalized SIPP* problems in [14]. Based on exchange scheme, an SDP relaxation method for solving SIPP problems was proposed in [33]. By representations of nonnegative polynomials in the univariate case, an SDP method was given in [35] for LSIPP problems (1.1) with S being closed intervals.

One of the difficulties in solving general SIP problems is the feasibility test of $\bar{u} \in \mathbb{R}^m$ which is equivalent to solve the problem of minimizing the constraint function with fixed \bar{u} over the index set. For SIPP or LSIPP problems, it is equivalent to a global polynomial optimization problem. Due to representations of nonnegative polynomials as sums of squares and the dual theory of moments, a hierarchy of SDP relaxations can be constructed to approximate polynomial optimization problems, see [13, 15, 19, 25] and the references therein. More precisely, we associate the set S with a so-called *quadratic module* which is a set of polynomials generated by $g_j(Y)$'s. If the *Archimedean* property (Definition 2.3) holds, then Putinar's Positivstellensatz [26] states that any polynomial positive over S belongs to the quadratic module. According to Schmüdgen's Positivstellensatz [29], if S is compact but the Archimedean property fails, we can replace in the above statement the quadratic module by the *preordering* generated by $g_j(Y)$'s. These representations of positive polynomials can be reduced to SDP feasibility problems. An SDP problem can be solved by interior-point method to a given accuracy in polynomial time [32, 34]. For this reason, it inspires us to investigate an SDP relaxation approach for solving the LSIPP problem (1.1).

Contribution. The main contribution of this paper is the following.

(i) We first consider the case when S is compact. By adding a redundant inequality in the description of S , we always assume that the Archimedean property is satisfied. When the Slater condition (Definition 2.1) holds for (1.1) which is commonly assumed in the literature on SIP, we show that the constraint that $a(y)^T x + b(y) \geq 0$ on S can be replaced by that the polynomial $a(Y)^T u + b(Y)$ belongs to the quadratic module associated with S . Then, a convergent hierarchy of SDP relaxations is constructed for (1.1). Thus, a decreasing sequence of upper bounds convergent to p^* can be computed. The practical importance of these upper bounds is that by combining the lower bounds of p^* gained by, for instance, discretization methods, a desired ε -optimal solution of (1.1) can be obtained.

(ii) We say that the finite convergence of the proposed SDP relaxations occurs if the optimal value of the SDP relaxation of some finite order equals p^* . We prove that a rank condition on the dual moment matrices of the SDP relaxations can be

used for certifying the finite convergence. We point out that this is only a sufficient condition which means that it might not hold when the finite convergence happens.

(iii) When S is noncompact, the Archimedean property fails. We use the homogenization technique to equivalently convert (1.1) into compact case under some generic assumption. Here, the generality means that the assumption holds if the coefficients of the polynomials g_j in the description of S are in general (see Remark 3.17). Consequently, a hierarchy of SDP relaxations of (1.1) with noncompact index set S is derived.

(iv) We apply this relaxation approach to the special LSIPP problem reformulated from a polynomial optimization problem. When the feasible set of the polynomial optimization problem is compact, we get the classic Lasserre's SDP relaxation method. When the feasible set is noncompact, a new and efficient SDP relaxation approach is obtained for solving the class of polynomial optimization problems whose objective polynomials are stably bounded from below on noncompact feasible sets. Note that the classic Lasserre's SDP relaxation method might fail for this class of problems, see Example 4.6.

This paper is organized as follows. We introduce some notation and preliminaries in Section 2. Depending on whether S is compact or not, two SDP relaxation methods and a stopping criterion are proposed in Section 3 for (1.1). In Section 4, we consider the application of the proposed relaxation approach to the special LSIPP problem reformulated from a polynomial optimization problem. A conclusion is made in Section 5.

2. NOTATION AND PRELIMINARIES

Here is some notation used in this paper. The symbol \mathbb{N} (resp., \mathbb{R}) denotes the set of nonnegative integers (resp., real numbers). For any $t \in \mathbb{R}$, $\lceil t \rceil$ denotes the smallest integer that is not smaller than t . For $y = (y_1, \dots, y_n) \in \mathbb{R}^n$, $\|y\|_2$ denotes the standard Euclidean norm of y . For $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, $\|\alpha\|_1 = \alpha_1 + \dots + \alpha_n$. For $k \in \mathbb{N}$, denote $\mathbb{N}_k^n = \{\alpha \in \mathbb{N}^n \mid \|\alpha\|_1 \leq k\}$. For $y \in \mathbb{R}^n$ and $\alpha \in \mathbb{N}^n$, y^α denotes $y_1^{\alpha_1} \dots y_n^{\alpha_n}$. $\mathbb{R}[Y] = \mathbb{R}[Y_1, \dots, Y_n]$ denotes the ring of polynomials in (Y_1, \dots, Y_n) with real coefficients. For $k \in \mathbb{N}$, denote by $\mathbb{R}[Y]_k$ the set of polynomials in $\mathbb{R}[Y]$ of total degree up to k . For a symmetric matrix W , $W \succeq 0$ ($\succ 0$) means that W is positive semidefinite (definite). For two symmetric matrices A, B of the same size, $\langle A, B \rangle$ denotes the inner product of A and B .

For any feasible point $x \in \mathbb{R}^m$ of (1.1), the *active index set* of x is

$$\{y \in S \mid a(y)^T x + b(y) = 0\}.$$

The *Haar dual* problem [1] of (1.1) is

$$(2.1) \quad (D) \quad \begin{cases} d^* := \sup & - \sum_{y \in S} \lambda_y b(y) \\ \text{s.t.} & \sum_{y \in S} \lambda_y a(y) = c, \\ & \lambda_y \geq 0, \forall y \in S, \end{cases}$$

where only finitely many dual variables λ_y , $y \in S$, take positive values.

Definition 2.1. We say that Slater condition holds for the problem (1.1) if there exists $\bar{x} \in \mathbb{R}^m$ such that $a(y)^T \bar{x} + b(y) > 0$ for all $y \in S$.

Proposition 2.2. [11, Theorem 6.9] *If S is compact and Slater condition holds for (1.1), then $p^* = d^*$ and d^* is attainable.*

Next we recall some background about *sums of squares* of polynomials and the dual theory of *moment matrices*. For any $f(Y) \in \mathbb{R}[Y]_k$, let \mathbf{f} denote its column vector of coefficients in the canonical monomial basis of $\mathbb{R}[Y]_k$. A polynomial $f(Y) \in \mathbb{R}[Y]$ is said to be a sum of squares of polynomials if it can be written as $f(Y) = \sum_{i=1}^t f_i(Y)^2$ for some $f_1(Y), \dots, f_t(Y) \in \mathbb{R}[Y]$. The symbol Σ^2 denotes the set of polynomials that are sums of squares.

Let $G := \{g_1, \dots, g_s\}$ be the set of polynomials that defines the semialgebraic set S (1.2). We denote

$$\mathcal{Q}(G) := \left\{ \sum_{j=0}^s \sigma_j g_j \mid g_0 = 1, \sigma_j \in \Sigma^2, j = 0, 1, \dots, s \right\}$$

as the *quadratic module* generated by G and its k -th *quadratic module*

$$\mathcal{Q}_k(G) := \left\{ \sum_{j=0}^s \sigma_j g_j \mid g_0 = 1, \sigma_j \in \Sigma^2, \deg(\sigma_j g_j) \leq 2k, j = 0, 1, \dots, s \right\}.$$

It is clear that if $f \in \mathcal{Q}(G)$, then $f(y) \geq 0$ for any $y \in S$. However, the converse is not necessarily true, see Example 3.12. Note that checking $f \in \mathcal{Q}_k(G)$ for a fixed $k \in \mathbb{N}$ is an SDP feasibility problem [12, 25].

For $k \in \mathbb{N}$, denote $s(k) := \binom{n+k}{n}$. A sequence of real numbers $z := (z_\alpha)_{\alpha \in \mathbb{N}_{2k}^n} \in \mathbb{R}^{s(2k)}$ whose elements are indexed by n -tuples $\alpha \in \mathbb{N}_{2k}^n$ is called a *truncated moment sequence* up to order $2k$. We say that $z \in \mathbb{R}^{s(2k)}$ has a representing measure μ if

$$z_\alpha = \int Y^\alpha d\mu(y), \quad \forall \alpha \in \mathbb{N}_{2k}^n.$$

The associated k -th *moment matrix* is the matrix $M_k(z)$ indexed by \mathbb{N}_k^n , with (α, β) -th entry $z_{\alpha+\beta}$ for $\alpha, \beta \in \mathbb{N}_k^n$. Given a polynomial $f(Y) = \sum_{\alpha} f_{\alpha} Y^{\alpha}$, for $k \geq d_f := \lceil \deg(f)/2 \rceil$, the $(k - d_f)$ -th *localizing moment matrix* $M_{k-d_f}(fz)$ is defined as the moment matrix of the *shifted vector* $((fz)_{\alpha})_{\alpha \in \mathbb{N}_{2(k-d_f)}^n}$ with $(fz)_{\alpha} = \sum_{\beta} f_{\beta} z_{\alpha+\beta}$. \mathcal{M}_{2k} denotes the space of all truncated moment sequences with order at most $2k$. For any $z \in \mathcal{M}_{2k}$, the Riesz functional \mathcal{L}_z on $\mathbb{R}[Y]_{2k}$ is defined by

$$\mathcal{L}_z \left(\sum_{\alpha} q_{\alpha} Y_1^{\alpha_1} \cdots Y_n^{\alpha_n} \right) := \sum_{\alpha} q_{\alpha} z_{\alpha}, \quad \forall q(Y) \in \mathbb{R}[Y]_{2k}.$$

From the definition of the localizing moment matrix $M_{k-d_f}(fz)$, it is easy to check that

$$(2.2) \quad \mathbf{q}^T M_{k-d_f}(fz) \mathbf{q} = \mathcal{L}_z(f(Y)q(Y)^2), \quad \forall q(Y) \in \mathbb{R}[Y]_{k-d_f}.$$

Let $d_j := \lceil \deg(g_j)/2 \rceil$ for each $j = 1, \dots, s$. For any $v \in S$, let $\zeta_{2k,v} := [v^{\alpha}]_{\alpha \in \mathbb{N}_{2k}^n}$ be the *Zeta vector* of v up to degree $2k$, i.e.,

$$\zeta_{2k,v} = [1 \quad v_1 \quad \cdots \quad v_n \quad v_1^2 \quad v_1 v_2 \quad \cdots \quad v_n^{2k}].$$

Then, $M_k(\zeta_{2k,v}) \succeq 0$ and $M_{k-d_j}(g_j \zeta_{2k,v}) \succeq 0$ for $j = 1, \dots, s$. In fact, let $g_0 = 1$, then for each $j = 0, 1, \dots, s$,

$$\mathbf{q}^T M_{k-d_j}(g_j \zeta_{2k,v}) \mathbf{q} = \mathcal{L}_{\zeta_{2k,v}}(g_j(Y)q(Y)^2) = g_j(v)q(v)^2 \geq 0, \quad \forall q(Y) \in \mathbb{R}[Y]_{k-d_j}.$$

Definition 2.3. We say that $\mathcal{Q}(G)$ is Archimedean if there exists $\psi \in \mathcal{Q}(G)$ such that the inequality $\psi(y) \geq 0$ defines a compact set in \mathbb{R}^n .

Note that the Archimedean property implies that S is compact but the converse is not necessarily true. However, for any compact set S we can always force the associated quadratic module to be Archimedean by adding a redundant constraint $M - \|y\|_2^2 \geq 0$ in the description of S for sufficiently large M .

Theorem 2.4. [26, Putinar's Positivstellensatz] Suppose that $\mathcal{Q}(G)$ is Archimedean.

- (i) If a polynomial $p \in \mathbb{R}[Y]$ is positive on S , then $p \in \mathcal{Q}_k(G)$ for some $k \in \mathbb{N}$;
- (ii) If $M_k(z) \succeq 0$ and $M_k(g_j z) \succeq 0$ for all $j = 1, \dots, s$, and all $k = 0, 1, \dots$, then z has a representing measure μ with support contained in S .

3. SDP RELAXATIONS OF LSIPP

In this section, depending on whether the index set S is compact or not, we shall construct two hierarchies of SDP relaxations and provide a sufficient stopping criterion when the finite convergence occurs for solving the LSIPP problem (1.1).

3.1. Compact case. We assume that S in (1.1) is compact.

3.1.1. SDP relaxations of compact LSIPP problems. For a given feasible point $x \in \mathbb{R}^m$ of the LSIPP problem (1.1), the constraint requires that the polynomial $a(Y)^T x + b(Y) \in \mathbb{R}[Y]$ is nonnegative on S . Since every polynomial in the quadratic module $\mathcal{Q}(G)$ of S generated by G is nonnegative on S , we can relax the problem (1.1) as follows

$$(3.1) \quad p^{\text{sos}} := \inf_{x \in \mathbb{R}^m} c^T x \quad \text{s.t.} \quad a(Y)^T x + b(Y) \in \mathcal{Q}(G).$$

Clearly, any feasible point of (3.1) is also feasible for (1.1). Hence, we have $p^{\text{sos}} \geq p^*$.

Theorem 3.1. If $\mathcal{Q}(G)$ is Archimedean and Slater condition holds for the LSIPP problem (1.1), then $p^{\text{sos}} = p^*$.

Proof. Fix an $\varepsilon > 0$ and a feasible $\bar{x} \in \mathbb{R}^m$ of (1.1) such that $a(y)^T \bar{x} + b(y) > 0$ for all $y \in S$. Since S is compact and (1.1) is linear, we have $c^T \bar{x} - p^* > 0$. Hence, we can fix another feasible point $x' \in \mathbb{R}^m$ of (1.1) such that $c^T \bar{x} > c^T x'$ and $c^T x' - p^* < \varepsilon/2$. Let

$$\delta := \frac{\varepsilon}{2c^T(\bar{x} - x')} > 0 \quad \text{and} \quad \hat{x} := (1 - \delta)x' + \delta\bar{x}.$$

Then

$$a(y)^T \hat{x} + b(y) = (1 - \delta)[a(y)^T x' + b(y)] + \delta[a(y)^T \bar{x} + b(y)] > 0, \quad \forall y \in S.$$

Since $\mathcal{Q}(G)$ is Archimedean, $a(Y)^T \hat{x} + b(Y) \in \mathcal{Q}(G)$ by Putinar's Positivstellensatz. That is, \hat{x} is feasible for both (1.1) and (3.1). We have

$$\begin{aligned} p^{\text{sos}} - p^* &\leq c^T \hat{x} - p^* \\ &= (1 - \delta)c^T x' + \delta c^T \bar{x} - p^* \\ &= (c^T x' - p^*) + \delta c^T(\bar{x} - x') \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary and $p^{\text{sos}} \geq p^*$, we conclude that $p^{\text{sos}} = p^*$. \square

Note that we do not require that p^* is attainable in the above proof. Define

$$(3.2) \quad \begin{aligned} d_j &:= \lceil \deg(g_j)/2 \rceil, \quad d_S := \max\{1, d_1, \dots, d_s\}, \\ d_P &:= \max\{d_S, \lceil \deg(a_1)/2 \rceil, \dots, \lceil \deg(a_m)/2 \rceil, \lceil \deg(b)/2 \rceil\}. \end{aligned}$$

For $k \geq d_P$, replacing $\mathcal{Q}(G)$ in (3.1) by its k -th truncation $\mathcal{Q}_k(G)$, we obtain

$$(3.3) \quad \begin{cases} p_k^{\text{sos}} := \inf_{x \in \mathbb{R}^m} c^T x \\ \text{s.t. } a(Y)^T x + b(Y) = \sum_{j=0}^s \sigma_j(Y) g_j(Y), \\ g_0 = 1, \sigma_j \in \Sigma^2, \deg(\sigma_j g_j) \leq 2k, \quad j = 0, \dots, s. \end{cases}$$

Now we reformulate (3.3) as an SDP problem. For any $t \in \mathbb{N}$, let $m_t(Y)$ be the column vector consisting of all the monomials in Y of degree up to t . Recall that $s(t) = \binom{n+t}{n}$ which is the cardinality of $m_t(Y)$. For each $j = 0, 1, \dots, s$, there exists a positive semidefinite matrix $Z_j \in \mathbb{R}^{s(k-d_j) \times s(k-d_j)}$ such that $\sigma_j(Y) = m_{k-d_j}(Y)^T \cdot Z_j \cdot m_{k-d_j}(Y)$. For each $\alpha \in \mathbb{N}_{2k}^n$, we can find a matrix $C_{j,\alpha} \in \mathbb{R}^{s(k-d_j) \times s(k-d_j)}$ such that the coefficient of $\sigma_j g_j$ equals $\langle Z_j, C_{j,\alpha} \rangle$ for each $j = 0, 1, \dots, s$. Let

$$b(Y) = \sum_{\alpha \in \mathbb{N}_{2k}^n} b_\alpha Y^\alpha \quad \text{and} \quad a_i(Y) = \sum_{\alpha \in \mathbb{N}_{2k}^n} a_{i,\alpha} Y^\alpha, \quad i = 1, \dots, m.$$

Then (3.3) can be written as the SDP problem

$$\begin{cases} p_k^{\text{sos}} = \inf_{Z_j \succeq 0, x \in \mathbb{R}^m} c^T x \\ \text{s.t. } \sum_{i=1}^m x_i a_{i,\alpha} + b_\alpha = \sum_{j=0}^s \langle Z_j, C_{j,\alpha} \rangle, \quad \forall \alpha \in \mathbb{N}_{2k}^n. \end{cases}$$

It follows that

Theorem 3.2. *If $\mathcal{Q}(G)$ is Archimedean and Slater condition holds for the LSIPP problem (1.1), then p_k^{sos} decreasingly converges to p^* as $k \rightarrow \infty$.*

Next we derive the dual problem of (3.3). Let $z = (z_\alpha)_{\alpha \in \mathbb{N}_{2k}^n} \in \mathbb{R}^{s(2k)}$ be a truncated moment sequence up to order $2k$. Then it is easy to check that the moment matrix $M_k(z)$ and the localizing moment matrix $M_{k-d_j}(g_j z)$ can be represented as

$$M_k(z) = \sum_{\alpha \in \mathbb{N}_{2k}^n} z_\alpha C_{0,\alpha} \quad \text{and} \quad M_{k-d_j}(g_j z) = \sum_{\alpha \in \mathbb{N}_{2k}^n} z_\alpha C_{j,\alpha}, \quad j = 1, \dots, s.$$

Then, the Lagrangian dual of the SDP problem (3.3) is the moment problem

$$(3.4) \quad \begin{cases} p_k^{\text{mom}} := \sup_{z \in \mathbb{R}^{s(2k)}} - \sum_{\alpha \in \mathbb{N}_{2k}^n} b_\alpha z_\alpha \\ \text{s.t. } \sum_{\alpha \in \mathbb{N}_{2k}^n} a_{i,\alpha} z_\alpha = c_i, \quad i = 1, \dots, m, \\ M_k(z) \succeq 0, \quad M_{k-d_j}(g_j z) \succeq 0, \quad j = 1, \dots, s. \end{cases}$$

By the ‘weak duality’, we have $p_k^{\text{mom}} \leq p_k^{\text{sos}}$.

Theorem 3.3. *If $\mathcal{Q}(G)$ is Archimedean and Slater condition holds for the LSIPP problem (1.1), then p_k^{mom} decreasingly converges to p^* as $k \rightarrow \infty$.*

Proof. By the ‘weak duality’ and Theorem 3.2, it suffices to prove that $p_k^{\text{mom}} \geq p^*$ for each $k \geq d_P$. Consider the Haar dual problem (2.1) of (1.1). Since S is compact and Slater condition holds for (1.1), by Proposition 2.2, $p^* = d^*$ and d^* is attainable. Denote $(\lambda_y^*)_{y \in S}$ as an optimizer of d^* and S^* as the finite subset of S such that $\lambda_y^* > 0$ for every $y \in S^*$. Let $\bar{z} = \sum_{y \in S^*} \lambda_y^* \zeta_{2k,y}$ where $\zeta_{2k,y}$ is the Zeta vector of y up to degree $2k$. Clearly, \bar{z} is feasible for (3.4) and then $p_k^{\text{mom}} \geq -\sum_{\alpha} b_{\alpha} \bar{z}_{\alpha} = d^* = p^*$. Hence, we obtain that $p_k^{\text{mom}} \downarrow p^*$ as $k \rightarrow \infty$. \square

Remark 3.4. Since S is compact, we can define another dual problem [11, 16, 30] of (1.1)

$$(3.5) \quad \begin{cases} \sup_{\mu \in M^+(S)} & - \int_S b(y) d\mu(y) \\ \text{s.t.} & \int_S a_i(y) d\mu(y) = c_i, \quad i = 1, \dots, m, \end{cases}$$

where $M^+(S)$ is the space of all nonnegative bounded regular Borel measure on S . Since S is compact, as shown in [28], the dual problems (2.1) and (3.5) have the same optimal value. By Putinar’s Positivstellensatz (part (ii) of Theorem 2.4), a moment sequence z has a representing measure μ with support contained in S if

$$M_k(z) \succeq 0, \quad M_k(g_j z) \succeq 0, \quad j = 1, \dots, s, \quad k = 0, 1, \dots$$

Therefore, (3.4) can be regarded as the k -th SDP relaxation of (3.5). \square

3.1.2. Optimality certificate. By the ‘weak duality’, the moment hierarchy (3.4) is tighter than (3.3). Moreover, (3.4) can be easily implemented and solved by the software GloptiPoly [10] developed by Henrion and Lasserre. We now propose a stopping criterion for (3.4) when the finite convergence occurs. Recall the notation in (3.2). Let $k \geq d_P$.

Condition 3.5. For an optimizer z^* of the k -th SDP relaxation (3.4), the following condition :

$$\exists t \in \mathbb{N} \quad \text{s.t.} \quad d_P \leq t \leq k \quad \text{and} \quad \text{rank} M_{t-d_S}(z^*) = \text{rank} M_t(z^*)$$

holds.

This condition can be used for certifying the finite convergence of Lasserre’s SDP relaxations [12] of polynomial optimization problems [13, 15, 21].

Theorem 3.6. Suppose that $\mathcal{Q}(G)$ is Archimedean and Slater condition holds for the LSIPP problem (1.1). If Condition 3.5 holds for an optimizer z^* of the k -th SDP relaxation (3.4), then $p_k^{\text{mom}} = p^*$.

Proof. By Theorem 3.3, it suffices to show that $p_k^{\text{mom}} \leq p^*$. Let $r = \text{rank} M_t(z^*)$. By [2, Theorem 1.1], z^* has a unique r -atomic measure supported on S , i.e., there exist r positive real numbers $\lambda_1, \dots, \lambda_r$ and r distinct points $v_1, \dots, v_r \in S$ such that

$$z^* = \lambda_1 \zeta_{2k,v_1} + \dots + \lambda_r \zeta_{2k,v_r},$$

where ζ_{2k,v_i} is the Zeta vector of v_i up to degree $2k$. By (3.4), we have

$$p_k^{\text{mom}} = - \sum_{i=1}^r \lambda_i b(v_i) \quad \text{and} \quad c = \sum_{i=1}^r \lambda_i a(v_i).$$

For any feasible point x of (1.1), we have

$$(3.6) \quad c^T x = \sum_{i=1}^r \lambda_i a(v_i)^T x \geq - \sum_{i=1}^r \lambda_i b(v_i) = p_k^{\text{mom}}.$$

The inequality is due to the feasibility of x . Hence, we have $p_k^{\text{mom}} \leq p^*$. \square

Remark 3.7. (i) The extraction procedure of the r points v_i in the above proof can be found in [9] and has been implemented in GloptiPoly. It can be inferred from (3.6) that the points v_1, \dots, v_r belong to the active index set of *every* minimizer x^* of the LSIPP problem (1.1).

(ii) Note that Condition 3.5 is only a sufficient condition which means that it might not hold when the finite convergence happens. Indeed, we will see in Section 4 that when applying (3.3) and (3.4) to the LSIPP problems reformulated from polynomial optimization problems, we can get their classic Lasserre's SDP relaxations whose finite convergence can be certified by Condition 3.5. Therefore, [15, Example 6.24] shows that Condition 3.5 is only sufficient not necessary. \square

3.1.3. Numerical experiments. The following are some testing examples which are implemented with MATLAB R2013a with two 3.60G cores and 12G RAM. We use GloptiPoly to manipulate the moment relaxation (3.4) and call the SDP solver SeDuMi [31] in GloptiPoly to solve the resulting SDP problems. The desired accuracy in SeDuMi is set to 10^{-8} . For checking Condition 3.5 and extracting the r points in the proof of Theorem 3.6, the singular value decomposition is used in GloptiPoly with accuracy set to 10^{-3} . The consumed computer time is calculated as the total time of all relaxations from the order $k = d_P$ to the order when Condition 3.5 is satisfied. The results show that Condition 3.5 holds for almost all examples.

Example 3.8. Consider the following optimization problem

$$\begin{aligned} \inf_{x \in \mathbb{R}^7} \quad & \sum_{i=1}^7 \frac{x_i}{i} \\ \text{s.t.} \quad & \sum_{i=1}^7 x_i y^{i-1} + \sum_{i=0}^4 y^{2i} \geq 0, \quad \forall y \in [0, 1]. \end{aligned}$$

This problem was studied in [4] and has an optimal solution of -1.78688 . We have $d_P = 4$. Computing the relaxations (3.4) with GloptiPoly, Condition 3.5 holds at the order $k = 4$ with $p_4^{\text{mom}} = -1.7869$. By Theorem 3.6, the finite convergence occurs and p_4^{mom} equals the optimal value. Here is a small numerical error due to the computations with finite precision in Matlab. The consumed computer time is 0.13 seconds. \square

Example 3.9. Consider the optimization problem

$$\begin{aligned} \inf_{x \in \mathbb{R}^2} \quad & x_2 \\ \text{s.t.} \quad & x_1 y_1 + x_2 - y_2 \geq 0, \quad \forall y \in S, \end{aligned}$$

where

$$S := \{y \in \mathbb{R}^2 \mid (y_1 + 5y_2)y_1^2 - (y_1^2 + y_2^2)^2 \geq 0\}$$

which is the gray region in Figure 1. Geometrically, the problem is to minimize the y_2 -intercept of the line $l_x(y) := x_1 y_1 + x_2 - y_2 = 0$ lying above S . From Figure 1, we can see that the optimum is reached when the line $l_x(y)$ is simultaneously

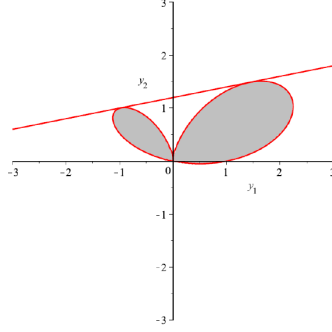


FIGURE 1. The semialgebraic set S in Example 3.9 and the optimal line $l_{x^*}(y) = 0$.

tangent at two different points of S . It can be checked that the optimal tangent line is

$$l_{x^*}(y) = \frac{1}{5}y_1 + \frac{125}{104} - y_2$$

and the two tangent points are

$$(3.7) \quad \begin{aligned} &\left(\frac{625}{2704} + \frac{1875}{2704}\sqrt{3}, \frac{3375}{2704} + \frac{375}{2704}\sqrt{3} \right) \approx (1.4322, 1.4884), \\ &\left(\frac{625}{2704} - \frac{1875}{2704}\sqrt{3}, \frac{3375}{2704} - \frac{375}{2704}\sqrt{3} \right) \approx (-0.9699, 1.0079). \end{aligned}$$

Hence, the optimum is $\frac{125}{104} \approx 1.2019$. We have $d_P = 2$. Compute the relaxations (3.4) with GloptiPoly. We get $p_2^{\text{mom}} = 1.2982$ and $p_3^{\text{mom}} = 1.2019$. Condition 3.5 holds at the order $k = 3$ and the extracted active index set, which corresponds to the set of points (3.7), consists of $(1.4321, 1.4883)$ and $(-0.9699, 1.0079)$. The consumed computer time is 0.24 seconds. \square

Example 3.10. Now we test the performance of the SDP relaxation (3.4) on some random LSIPP problems which are generated as follows.

Let $S = [-2, 2]^n$. Randomly pick m distinct points $v^{(1)}, \dots, v^{(m)}$ from S whose coordinates are integers drawn from the discrete uniform distribution on $[-2, 2]$. Let $a_1(Y), \dots, a_m(Y)$ be the Lagrange interpolation polynomials at these m points, i.e., for each $v^{(i)}$ and $j \neq i$, randomly choose an index $k_{i,j} \in \{1, \dots, n\}$ for which $v_{k_{i,j}}^{(i)} \neq v_{k_{i,j}}^{(j)}$ and define

$$a_i(Y) = \prod_{j \neq i} \frac{Y_{k_{i,j}} - v_{k_{i,j}}^{(j)}}{v_{k_{i,j}}^{(i)} - v_{k_{i,j}}^{(j)}}, \quad i = 1, \dots, m.$$

Then, we have $a_i(v^{(i)}) = 1$ and $a_i(v^{(j)}) = 0$ for each $j \neq i$. Recall that for any $t \in \mathbb{N}$, $m_t(Y)$ denotes the column vector consisting of all the monomials in Y of degree up to t and $s(t)$ denotes its cardinality. Let N be an $s(t) \times s(t)$ matrix containing random elements drawn from the standard uniform distribution on the open interval $(0, 1)$. Define $b(Y) = (N \cdot m_t(Y))^T (N \cdot m_t(Y)) + 1$. For each $i = 1, \dots, m$, also choose c_i from the standard uniform distribution on the open interval $(0, 1)$. We add an redundant inequality $\|Y\|_2^2 \leq 4n$ in the description of S to make the Archimedean property hold.

TABLE 1. Computational results for random LSIPP problems

No.	(m, n, t)	<i>inst.</i>	<i>certi.</i>	<i>order</i> (min, max)		<i>time</i> (min, max)	
1	(5, 3, 2)	20	20	2	3	0.28s	0.65s
2	(6, 3, 2)	20	20	3	4	0.38s	0.98s
3	(7, 3, 2)	20	20	3	4	0.41s	1.11s
4	(8, 3, 2)	20	20	4	4	0.69s	0.83s
5	(9, 3, 2)	20	20	4	5	0.78s	2.83s
6	(10, 3, 2)	20	20	5	6	1.88s	7.61s
7	(11, 3, 2)	20	20	5	6	1.90s	8.69s
8	(12, 3, 2)	20	15	6	6	5.63s	32.93s
9	(5, 4, 2)	20	20	2	3	0.38s	0.98s
10	(5, 5, 2)	20	20	2	3	0.51s	2.38s
11	(5, 6, 2)	20	20	2	3	0.76s	7.41s
12	(5, 7, 2)	20	20	2	3	1.22s	30.42s
13	(5, 8, 2)	20	20	2	3	2.11s	254.31s
14	(5, 4, 3)	20	20	3	3	0.92s	1.10s
15	(5, 4, 4)	20	20	4	4	4.35s	5.29s
16	(5, 4, 5)	20	20	5	5	23.50s	30.07s
17	(5, 4, 6)	20	20	6	6	144.16s	185.98s

For the class of random LSIPP problems constructed above, we can see that the Slater condition holds if we let $\bar{x} = 0$. Moreover, for each $i = 1, \dots, m$, we have $x_i \geq -b(v_i) - 1$ for every feasible point $x \in \mathbb{R}^m$. Therefore, the optimum $p^* > -\infty$. Hence, all assumptions needed for the convergence of the SDP relaxations (3.3) and (3.4) are satisfied.

We test several groups of the above random LSIPP problems with the SDP relaxations (3.4) from the order $k = d_P$ to the order $k = 8$. The results are shown in Table 1 where the *inst.* column denotes the number of randomly generated instances and the *certi.* column denotes the number of instances where certified finite convergence occurs, i.e., Condition 3.5 holds at some order. Among all instances of each group, the min (max) *order* column shows the minimal (maximal) order of relaxations when Condition 3.5 is satisfied and the min (max) *time* column shows the minimal (maximal) consumed computer time. \square

Remark 3.11. As a summary, compared with the existing methods in the literature applied to the LSIPP problems (1.1), our SDP relaxations approach of (3.3) and (3.4) has the following properties: (a) A decreasing sequence of upper bounds of p^* can be computed by solving a sequence of SDP problems; (b) The global convergence is guaranteed under some mild assumptions (Theorem 3.2 and 3.3). Hence, by combining the lower bounds of p^* gained by, for instance, discretization methods, a desired ε -optimal solution of (1.1) can be obtained. (c) A checkable sufficient condition is available for certifying the finite convergence of the SDP relaxations when it occurs (Theorem 3.6); (d) Obviously, the sizes of the resulting SDP problems grow exponentially as the number of parameters and the order increase. Thus, our SDP relaxation method is more suitable for small or medium size LSIPP problems.

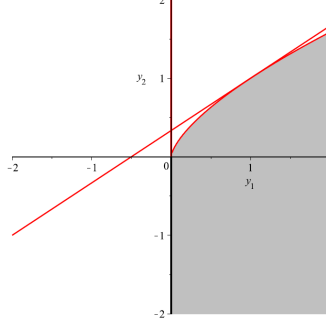


FIGURE 2. The semialgebraic set S in Example 3.12 and the tangent line $l_{3/2}(y) = 0$.

3.2. Noncompact case. In this section, we consider the LSIPP problem (1.1) with noncompact index set S . Since the Archimedean property is violated in this case, the optima of the SDP relaxations (3.3) and (3.4) might not converge to p^* . We illustrate it by the following example.

Example 3.12. Consider the LSIPP problem

$$(3.8) \quad \inf_{x \in \mathbb{R}} -\frac{x}{2} \quad \text{s.t.} \quad (1 - 3y_2)x + 3y_1 \geq 0, \quad \forall y \in S,$$

where

$$S := \{y \in \mathbb{R}^2 \mid y_1 \geq 0, y_1^2 - y_2^3 \geq 0\}$$

which is the gray shadow below the right half of the cusp as shown in Figure 2. Since $(0, 0) \in S$, a feasible x must be nonnegative. Clearly, $x = 0$ is a feasible point. When $x > 0$, the constraint in (3.8) means that S lies in the half plane defined by $l_x(y) := y_2 - y_1/x - 1/3 \leq 0$. Hence, the optimum is attained when the line $l_x(y) = 0$ is tangent to S . It is easy to check that the feasible set is $[0, 3/2]$ and the optimum is $-3/4$. The tangent line $l_{3/2}(y) = 0$ is shown red in Figure 2. The Slater condition holds for any point $x \in (0, 3/2)$.

Let $G := \{Y_1, Y_1^2 - Y_2^3\}$. Then, $\mathcal{Q}(G)$ is not Archimedean since S is noncompact. For any $k \in \mathbb{N}$, we know from [8, Example 2.10] that $(1 - 3Y_2)x + 3Y_1 \in \mathcal{Q}_k(G)$ if and only if $x = 0$, i.e., $p_k^{\text{sos}} = 0$ for each $k \geq d_P$. Now we show that $p_k^{\text{mom}} = p_k^{\text{sos}}$ for each $k \geq d_P$. In fact, for the SDP relaxation (3.4) of the problem (3.8), let μ be a probability measure with uniform distribution in the following subset of S :

$$S_1 := \{(y_1, y_2) \in \mathbb{R}^2 \mid 1 \leq y_1 \leq 2, 0 \leq y_2 \leq 1\}$$

and $z^{(\mu)}$ be the truncated moment sequence with representing measure μ up to order $2k$. It can be verified that $z^{(\mu)}$ is a feasible point of (3.4) and its corresponding truncated moment matrix and localizing moment matrices are positive definite since S_1 has nonempty interior. Then $p_k^{\text{mom}} = p_k^{\text{sos}}$ follows by the conic duality theorem. Hence, both SDP relaxations (3.3) and (3.4) do not converge to the optimum. \square

In [33], we used the technique of homogenization to convert a general SIPP problem with noncompact index set into compact case. In the following, we apply this technique to (1.1).

For a polynomial $f(Y) \in \mathbb{R}[Y]$, denote its homogenization by $f^h(\tilde{Y}) \in \mathbb{R}[\tilde{Y}]$, where $\tilde{Y} = (Y_0, Y_1, \dots, Y_n)$, i.e., $f^h(\tilde{Y}) = Y_0^{D_f} f(Y/Y_0)$, $D_f = \deg(f)$. For the

basic semialgebraic set S in (1.1), define

$$(3.9) \quad \begin{aligned} \tilde{S}_{>} &:= \{\tilde{y} \in \mathbb{R}^{n+1} \mid g_1^h(\tilde{y}) \geq 0, \dots, g_s^h(\tilde{y}) \geq 0, y_0 > 0, \|\tilde{y}\|_2^2 = 1\}, \\ \tilde{S} &:= \{\tilde{y} \in \mathbb{R}^{n+1} \mid g_1^h(\tilde{y}) \geq 0, \dots, g_s^h(\tilde{y}) \geq 0, y_0 \geq 0, \|\tilde{y}\|_2^2 = 1\}. \end{aligned}$$

Proposition 3.13. [33, Proposition 4.2] *For any $f(Y) \in \mathbb{R}[Y]$, $f(y) \geq 0$ on S if and only if $f^h(\tilde{y}) \geq 0$ on $\text{closure}(\tilde{S}_{>})$.*

Define

$$\omega := \max\{\deg(a_1), \dots, \deg(a_m), \deg(b)\}.$$

We homogenize the polynomials $a_i(Y)$, $i = 1, \dots, m$, and $b(Y)$ to the *same* degree ω and still denote the resulting polynomials as $a_i^h(\tilde{Y})$ and $b^h(\tilde{Y})$ for simplicity. Denote $a^h(\tilde{Y}) = (a_1^h(\tilde{Y}), \dots, a_m^h(\tilde{Y}))$. It follows that the problem (1.1) is equivalent to

$$\begin{cases} \inf_{x \in \mathbb{R}^m} c^T x \\ \text{s.t. } a^h(\tilde{y})^T x + b^h(\tilde{y}) \geq 0, \quad \forall \tilde{y} \in \text{closure}(\tilde{S}_{>}). \end{cases}$$

Replacing $\text{closure}(\tilde{S}_{>})$ by the basic semialgebraic set \tilde{S} , we get the following problem

$$(3.10) \quad \begin{cases} \tilde{p}^* := \inf_{x \in \mathbb{R}^m} c^T x \\ \text{s.t. } a^h(\tilde{y})^T x + b^h(\tilde{y}) \geq 0, \quad \forall \tilde{y} \in \tilde{S}. \end{cases}$$

It is obvious that $\tilde{p}^* \geq p^*$ since $\text{closure}(\tilde{S}_{>}) \subseteq \tilde{S}$.

For any polynomial $f(Y) \in \mathbb{R}[Y]$, denote $\hat{f}(Y)$ as its homogeneous part of the highest degree. Define

$$(3.11) \quad \hat{S} := \{y \in \mathbb{R}^n \mid \hat{g}_1(y) \geq 0, \dots, \hat{g}_s(y) \geq 0, \|y\|_2^2 = 1\}.$$

Specially, denote $\hat{a}_i(Y)$, $i = 1, \dots, m$, and $\hat{b}(Y)$ as the homogeneous parts of $a_i(Y)$, $i = 1, \dots, m$, and $b(Y)$ of the *same* degree ω . Let $\hat{a}(Y) := (\hat{a}_1(Y), \dots, \hat{a}_m(Y))$.

Condition 3.14. *For any $\varepsilon > 0$, there exists a feasible point $x^{(\varepsilon)}$ of (1.1) such that*

$$c^T x^{(\varepsilon)} - p^* \leq \varepsilon \quad \text{and} \quad \hat{a}(y)^T x^{(\varepsilon)} + \hat{b}(y) \geq 0, \quad \forall y \in \hat{S}.$$

Theorem 3.15. $\tilde{p}^* = p^*$ if and only if Condition 3.14 holds for (1.1).

Proof. By Proposition 3.13 and the fact that $\tilde{S} \setminus \text{closure}(\tilde{S}_{>}) \subseteq \{0\} \times \hat{S}$, it is straightforward to verify the conclusion. \square

Definition 3.16. [22] *We say that S is closed at ∞ if $\text{closure}(\tilde{S}_{>}) = \tilde{S}$.*

Remark 3.17. Clearly, $\tilde{p}^* = p^*$ when S is closed at ∞ . Note that not every set S of form (1.2) is closed at ∞ even when it is compact [20, Example 5.2]. However, it is shown in [33, Theorem 4.10] that the closedness at ∞ is a *generic* property. Namely, if we consider the space of all coefficients of generators g_j 's of all possible sets S of form (1.2) in the canonical monomial basis of $\mathbb{R}[Y]_d$, coefficients of g_j 's of those sets S which are not closed at ∞ are in a Zariski closed set of the space. It follows that the problems (1.1) and (3.10) are equivalent *in general*. Note that $\tilde{S}_{>}$ depends only on S , while \tilde{S} depends not only on S but also on the choice of the inequalities $g_1(y) \geq 0, \dots, g_s(y) \geq 0$. In some cases, we can add some redundant inequalities in the description of S to force it to be closed at ∞ [8].

Next we construct the corresponding SDP relaxations (3.3) and (3.4) of the problem (3.10). Let

$$(3.12) \quad G^h := \{g_1^h, \dots, g_s^h, Y_0, \|\tilde{Y}\|_2^2 - 1, 1 - \|\tilde{Y}\|_2^2\}$$

and denote $\mathcal{Q}(G^h)$ as the quadratic module of \tilde{S} generated by G^h . Then (3.3) becomes

$$(3.13) \quad \begin{cases} \tilde{p}_k^{\text{sos}} := \inf_{x \in \mathbb{R}^m} c^T x \\ \text{s.t. } a^h(\tilde{Y})^T x + b^h(\tilde{Y}) \in \mathcal{Q}(G^h). \end{cases}$$

For $k \in \mathbb{N}$, denote $\tilde{s}(k) := \binom{n+k+1}{n+1}$. Let $z := (z_{\tilde{\alpha}})_{\tilde{\alpha} \in \mathbb{N}_{2k}^{n+1}} \in \mathbb{R}^{\tilde{s}(2k)}$ be a truncated moment sequence up to order $2k$ whose elements are indexed by $(n+1)$ -tuples $\tilde{\alpha} := (\alpha_0, \alpha_1, \dots, \alpha_n) \in \mathbb{N}_{2k}^{n+1}$. Let

$$b^h(\tilde{Y}) = \sum_{\tilde{\alpha} \in \mathbb{N}_{2k}^{n+1}} b_{\tilde{\alpha}}^h \tilde{Y}^{\tilde{\alpha}} \quad \text{and} \quad a_i^h(\tilde{Y}) = \sum_{\tilde{\alpha} \in \mathbb{N}_{2k}^{n+1}} a_{i,\tilde{\alpha}}^h \tilde{Y}^{\tilde{\alpha}}, \quad i = 1, \dots, m.$$

According to (3.4), the dual of (3.13) is

$$(3.14) \quad \begin{cases} \tilde{p}_k^{\text{mom}} := \sup_{z \in \mathbb{N}_{2k}^{n+1}} - \sum_{\tilde{\alpha} \in \mathbb{N}_{2k}^{n+1}} b_{\tilde{\alpha}}^h z_{\tilde{\alpha}} \\ \text{s.t. } \sum_{\tilde{\alpha} \in \mathbb{N}_{2k}^{n+1}} a_{i,\tilde{\alpha}}^h z_{\tilde{\alpha}} = c_i, \quad i = 1, \dots, m, \\ M_k(z) \succeq 0, \quad M_{k-d_j}(g_j^h z) \succeq 0, \quad j = 1, \dots, s, \\ M_{k-1}(Y_0 z) \succeq 0, \quad M_{k-1}((\|\tilde{Y}\|_2^2 - 1)z) = 0. \end{cases}$$

Condition 3.18. *There exists a point $\bar{x} \in \mathbb{R}^m$ of (1.1) such that $a(y)^T \bar{x} + b(y) > 0$ for all $y \in S$ and $\hat{a}(y)^T \bar{x} + \hat{b}(y) > 0$ for all $y \in \hat{S}$.*

Proposition 3.19. *Slater condition holds for (3.10) if and only if Condition 3.18 holds for (1.1).*

Proof. Suppose Condition 3.18 holds for (1.1) at \bar{x} . For any $\tilde{v} = (v_0, v) \in \tilde{S}$, we have $v \in \hat{S}$ if $v_0 = 0$ and $v/v_0 \in S$ otherwise. It is straightforward to verify that Slater condition also holds for (3.10) at \bar{x} .

Suppose that Slater condition holds for (3.10) at $\bar{x} \in \mathbb{R}^m$. For any point $v \in \mathbb{R}^n$, we have $(0, v) \in \tilde{S}$ if $v \in \hat{S}$ and $\left(\frac{1}{\sqrt{1+\|v\|_2^2}}, \frac{v}{\sqrt{1+\|v\|_2^2}}\right) \in \tilde{S}$ if $v \in S$. Then similarly, it implies that Condition 3.18 holds for (1.1) at \bar{x} . \square

Theorem 3.20. *If Condition 3.18 holds for (1.1), then both \tilde{p}_k^{sos} and \tilde{p}_k^{mom} decreasingly converge to \tilde{p}^* as $k \rightarrow \infty$. Moreover, they both converge to p^* if S is closed at ∞ or Condition 3.14 holds for (1.1).*

Proof. Since $\mathcal{Q}(G^h)$ is Archimedean, the conclusion follows by combining Theorem 3.2, 3.3, 3.15 and Proposition 3.19. \square

Example 3.12 continue. By definition, we have

$$\begin{aligned}\tilde{S}_{>} &= \{(y_0, y_1, y_2) \in \mathbb{R}^3 \mid y_1 \geq 0, y_0 y_1^2 - y_2^3 \geq 0, y_0 > 0, \|\tilde{y}\|_2^2 = 1\}, \\ \tilde{S} &= \{(y_0, y_1, y_2) \in \mathbb{R}^3 \mid y_1 \geq 0, y_0 y_1^2 - y_2^3 \geq 0, y_0 \geq 0, \|\tilde{y}\|_2^2 = 1\}, \\ \hat{S} &= \{(y_1, y_2) \in \mathbb{R}^2 \mid y_1 \geq 0, y_2 \leq 0, y_1^2 + y_2^2 = 1\}.\end{aligned}$$

After homogenization, the problem (3.8) is reformulated as

$$(3.15) \quad \inf_{x \in \mathbb{R}} -\frac{x}{2} \quad \text{s.t. } (y_0 - 3y_2)x + 3y_1 \geq 0, \forall \tilde{y} = (y_0, y_1, y_2) \in \tilde{S},$$

Note that S is closed at ∞ . In fact, for every $(0, v_1, v_2) \in \tilde{S} \setminus \tilde{S}_{>}$, let

$$v^{(\varepsilon)} := \left(\varepsilon, v_1, \sqrt[3]{\varepsilon v_1^2 + v_2^3} \right).$$

Then $\{v^{(\varepsilon)} / \|v^{(\varepsilon)}\|_2\}_{\varepsilon > 0} \subseteq \tilde{S}_{>}$ and $\lim_{\varepsilon \rightarrow 0} v^{(\varepsilon)} / \|v^{(\varepsilon)}\|_2 = (0, v_1, v_2)$. Hence, we have $\tilde{S} \setminus \tilde{S}_{>} \subseteq \text{closure}(\tilde{S}_{>})$ and S is closed at ∞ . It is easy to check that Condition 3.18 holds for (3.15) if we let $\bar{x} = 1$. Hence, the assumptions in Theorem 3.20 are satisfied. With GloptiPoly, we get the following numerical results: $\tilde{p}_2^{\text{mom}} = -1.2124 \times 10^{-8}$ and $\tilde{p}_3^{\text{mom}} = -0.7500$. Condition 3.5 is satisfied for $k = 3$ and we obtain the certified optimum -0.7500 . As noted in Remark 3.7, the extracted numerical active index set of the minimizer $x^* = 3/2$ is $(0.5773, 0.5774, 0.5774)$ which corresponds to $(1, 1) \in S$ where the line $l_{x^*}(y) = 0$ is tangent to S . \square

4. SPECIAL CASE: POLYNOMIAL OPTIMIZATION PROBLEMS

Consider the general polynomial optimization problem

$$(4.1) \quad \begin{cases} f^* := \inf_{y \in \mathbb{R}^n} f(y) \\ \text{s.t. } g_1(y) \geq 0, \dots, g_s(y) \geq 0. \end{cases}$$

Recall that the feasible set of (4.1) is denoted as S . We assume that $-\infty < f^* < \infty$. The problem (4.1) can be reformulated as an LSIPP problem

$$(4.2) \quad -f^* = \inf_{x \in \mathbb{R}} -x \quad \text{s.t. } f(y) - x \geq 0, \forall y \in S.$$

As we will see, by applying the SDP relaxation approach derived in Section 3 to the special LSIPP problem (4.2), we can obtain:

- (i) the classic Lasserre's SDP relaxation method [12] of (4.1) when S is compact, which can be expected from the way of reformulation (4.2) and relaxation (3.1);
- (ii) a new and efficient hierarchy of SDP relaxations of (4.1) when S is noncompact and f is *stably bounded from below* on S , which is a class of polynomial optimization problems studied in [18]. Note that the classic Lasserre's SDP relaxation method might fail for this kind of problems, see Example 4.6.

4.1. Compact case. We first assume that S is compact. In the special LSIPP (4.2), we have

$$m = 1, a(Y) = -1, b(Y) = f(Y) \text{ and } c = -1.$$

According to (3.3) and (3.4), by exchanging of ‘inf’ and ‘sup’, we obtain SDP relaxations of (4.1):

$$(4.3) \quad \begin{cases} f_k^{\text{sos}} := \sup_{x \in \mathbb{R}} x \\ \text{s.t. } f(Y) - x = \sum_{j=0}^s \sigma_j(Y) g_j(Y), \\ g_0 = 1, \sigma_j \in \Sigma^2, \deg(\sigma_j g_j) \leq 2k, \quad i = 0, \dots, s, \end{cases}$$

and

$$(4.4) \quad \begin{cases} f_k^{\text{mom}} := \inf_{z \in \mathbb{R}^{s(2k)}} \sum_{\alpha \in \mathbb{N}_{2k}^s} f_\alpha z_\alpha \\ \text{s.t. } z_{\mathbf{0}} = 1, \\ M_k(z) \succeq 0, \quad M_{k-d_j}(g_j z) \succeq 0, \quad j = 1, \dots, s, \end{cases}$$

where $z_{\mathbf{0}}$ denotes the element of z indexed by the n -tuple $(0, \dots, 0)$. They are just the classic Lasserre’s SDP relaxations of polynomial optimization problems [12]. Clearly, Slater condition holds for (4.2) if and only if f is bounded from below on S . Hence, by Theorem 3.2 and Theorem 3.3, when $\mathcal{Q}(G)$ is Archimedean, both f_k^{sos} and f_k^{mom} converge to f^* as $k \rightarrow \infty$, which has already been proved in [12, Theorem 4.2].

Note that by Remark 3.7 the points in the active index set of the minimizer of (4.2) extracted when Condition 3.5 is satisfied are just the global minimizers of (4.1).

4.2. Noncompact case. Now we consider the polynomial optimization problem (4.1) with noncompact feasible set S . After homogenization, the problem (4.2) becomes

$$(4.5) \quad \begin{cases} \tilde{f}^* := \sup_{x \in \mathbb{R}} x \\ \text{s.t. } f^h(\tilde{y}) - x y_0^{D_f} \geq 0, \quad \forall \tilde{y} \in \tilde{S}, \end{cases}$$

where $D_f = \deg(f)$. According to (3.3) and (3.4), we obtain a hierarchy of SDP relaxations of (4.5):

$$(4.6) \quad \begin{cases} \tilde{f}_k^{\text{sos}} := \sup_{x \in \mathbb{R}} x \\ \text{s.t. } f^h(\tilde{Y}) - x Y_0^{D_f} \in \mathcal{Q}_k(G^h), \end{cases}$$

where G^h is defined in (3.12). For $k \in \mathbb{N}$, denote $\tilde{s}(k) := \binom{n+k+1}{n+1}$. Let $z := (z_{\tilde{\alpha}})_{\tilde{\alpha} \in \mathbb{N}_{2k}^{n+1}} \in \mathbb{R}^{\tilde{s}(2k)}$ be a truncated moment sequence of degree $2k$. Denote $z_{D_f, \mathbf{0}}$ as the element of z indexed by the $(n+1)$ -tuple $(D_f, 0, \dots, 0)$. The dual problem of (4.6) is

$$(4.7) \quad \begin{cases} \tilde{f}_k^{\text{mom}} := \inf_{z \in \mathbb{R}^{\tilde{s}(2k)}} \sum_{\tilde{\alpha} \in \mathbb{N}_{2k}^{n+1}} f_{\tilde{\alpha}}^h z_{\tilde{\alpha}} \\ \text{s.t. } z_{D_f, \mathbf{0}} = 1, M_k(z) \succeq 0, M_{k-1}(Y_0 z) \succeq 0, \\ M_{k-d_j}(g_j^h z) \succeq 0, \quad j = 1, \dots, s, \\ M_{k-1}((\|\tilde{Y}\|_2^2 - 1)z) = 0. \end{cases}$$

Definition 4.1. [18] We say that f is stably bounded from below on S if f remains bounded from below on S for all sufficiently small perturbations of the coefficients of f, g_1, \dots, g_s .

Recall the notation $\hat{f}(Y)$ and \hat{S} defined in (3.11).

Proposition 4.2. The following conditions are equivalent:

- (i) f is stably bounded from below on S ;
- (ii) \hat{f} is strictly positive on \hat{S} ;
- (iii) Slater condition holds for (4.5).

Proof. By [18, Theorem 5.1 and 5.3], (i) f is stably bounded from below on S if and only if the function $\max\{-\hat{g}_1, \dots, -\hat{g}_s, \hat{f}\}$ is strictly positive on the unit sphere, which is equivalent to (ii). The equivalence of (ii) and (iii) follows from Proposition 3.19. \square

Theorem 4.3. If either of the conditions in Proposition 4.2 holds, then: (a) $\tilde{f}^* = f^*$; (b) both \tilde{f}_k^{sos} and \tilde{f}_k^{mom} converge to f^* as $k \rightarrow \infty$.

Proof. If condition (ii) in Proposition 4.2 is satisfied, then Condition 3.14 holds for (4.2) and hence $\tilde{f}^* = f^*$ by Theorem 3.15. The conclusion follows by combining Theorem 3.2, 3.3, 3.15 and Proposition 4.2. \square

Remark 4.4. Suppose that one of the conditions in Proposition 4.2 holds. When solving (4.7), for any point $(v_0, v) \in \mathbb{R}^{n+1}$ with $v_0 \neq 0$ in the extracted active index set when Condition 3.5 is satisfied, it is easy to check that the point v/v_0 is a minimizer of $f(Y)$ on S .

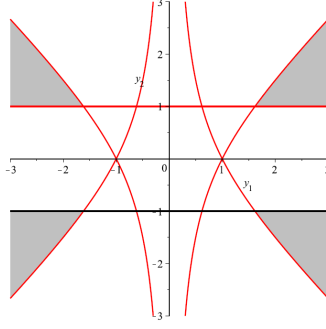
Remark 4.5. For solving the class of polynomial optimization problems (4.1) with f being stably bounded from below on a noncompact feasible set S , Marshall [18, Corollary 5.4] proposed the following method. For $\alpha \in \mathbb{N}^n$, let f_α and $g_{j,\alpha}$ denote the coefficients of Y^α in f and g_j , $j = 1, \dots, s$. Fix a lower bound $\varepsilon > 0$ of $\max\{-\hat{g}_1, \dots, -\hat{g}_s, \hat{f}\}$ on the unit sphere. Normalize so that $0 \in S$ and $f(0) = 0$. Then minimizing f on S is equivalent to minimizing f on the compact set $S \cap \{y \in \mathbb{R}^n \mid \|y\|_2^2 \leq \rho_\varepsilon^2\}$ where

$$\rho_\varepsilon = \max \left\{ 1, \sum_{\|\alpha\|_1 < \deg f} |f_\alpha|/\varepsilon, \sum_{\|\alpha\|_1 < \deg g_j} |g_{j,\alpha}|/\varepsilon : j = 1, \dots, s \right\}.$$

As pointed in [18, Notes 5.2], the lower bound ε can be computed by solving $s+1$ polynomial optimization problems on compact semialgebraic sets and hence there are obvious problems with this if s is too large. As we have seen, we can instead equivalently reformulate (4.1) as (4.5) and solve the *single* optimization problem by SDP relaxations (4.6) and (4.7). \square

Example 4.6. Consider the following polynomial optimization problem

$$(4.8) \quad \begin{cases} \inf_{y \in \mathbb{R}^2} & y_1^2 + y_2^2 \\ \text{s.t.} & y_2^2 - 1 \geq 0, \\ & y_1^2 - My_1y_2 - 1 \geq 0, \\ & y_1^2 + My_1y_2 - 1 \geq 0, \end{cases}$$

FIGURE 3. The semialgebraic set S in Example 4.6

where M is a positive constant. It was shown in [3, 17, 22] that the global minimizers and global minimum are

$$\left(\pm \frac{M + \sqrt{M^2 + 4}}{2}, \pm 1 \right) \quad \text{and} \quad 2 + \frac{M(M + \sqrt{M^2 + 4})}{2}.$$

We consider the case when $M = 1$. The feasible set S is depicted in gray in Figure 3. The global minimizers are $\left(\pm \frac{1+\sqrt{5}}{2}, \pm 1 \right) \approx (\pm 1.618, \pm 1)$ and its global minimum is $2 + \frac{(1+\sqrt{5})}{2} \approx 3.618$. Because S is noncompact, by the argument in [3], the classic Lasserre's SDP relaxations (4.3) of (4.8) can only provide lower bound $f_k^{\text{sos}} = 2$ no matter how large the order k is. Since S has nonempty interior and the relaxation (4.3) of (4.8) is feasible, by [12, Theorem 4.2], f_k^{mom} equals f_k^{sos} for each k and therefore can not converge to the optimum as $k \rightarrow \infty$, either.

Obviously, the condition (ii) in Proposition 4.2 holds. We compute the relaxations (4.7) with GloptiPoly. For $k = 3$, Condition 3.5 is satisfied and we get the numerically certified optimum $f_3^{\text{mom}} = 3.6180$. The extracted active index set is $\{(0.4653, \pm 0.7529, \pm 0.4653)\}$ which, by Remark 4.4, provides the global minimizers $(\pm 1.6181, \pm 1)$. \square

5. CONCLUSION

In this paper, we study a subclass of semi-infinite programming problems whose constraint functions are polynomials in parameters and index sets are basic semi-algebraic sets (LSIPP problems). When the index set of an LSIPP problem is compact, a convergent hierarchy of SDP relaxations is constructed based on Putinar's Positivstellensatz. We extend this approach to the case when the index set is noncompact by the technique of homogenization. Applying our method to the LSIPP problem reformulated from a polynomial optimization problem, we obtain a new hierarchy of SDP relaxations for solving the class of polynomial optimization problems whose objective polynomials are stably bounded from below on noncompact feasible sets.

Acknowledgments: The author is supported by the Chinese National Natural Science Foundation under grants 11401074. This work is supported by Fields Institute July-December 2015 Thematic Program on Computer Algebra in Toronto, Canada.

REFERENCES

- [1] A. Charnes, W. W. Cooper, and K. Kortanek. Duality in semi-infinite programs and some works of Haar and Carathéodory. *Management Science*, 9(2):209–228, 1963.
- [2] R. E. Curto and L. A. Fialkow. Truncated K -moment problems in several variables. *Journal of Operator Theory*, 54(1):189–226, 2005.
- [3] J. Demmel, J. Nie, and V. Powers. Representations of positive polynomials on noncompact semialgebraic sets via KKT ideals. *Journal of Pure and Applied Algebra*, 209(1):189 – 200, 2007.
- [4] M. C. Ferris and A. B. Philpott. An interior point algorithm for semi-infinite linear programming. *Mathematical Programming*, 43(1):257–276, 1989.
- [5] M. A. Goberna. Linear semi-infinite optimization: Recent advances. In V. Jeyakumar and A. Rubinov, editors, *Continuous Optimization*, volume 99 of *Applied Optimization*, pages 3–22. Springer, US, 2005.
- [6] M. A. Goberna and M. A. López. A comprehensive survey of linear semi-infinite optimization theory. In R. Reemtsen and J.-J. Rckmann, editors, *Semi-Infinite Programming*, volume 25 of *Nonconvex Optimization and Its Applications*, pages 3–27. Springer US, 1998.
- [7] M. A. Goberna and M. A. López. *Linear semi-infinite optimization*. John Wiley & Sons, Chichester, 1998.
- [8] F. Guo, C. Wang, and L. Zhi. Semidefinite representations of noncompact convex sets. *SIAM Journal on Optimization*, 25(1):377–395, 2015.
- [9] D. Henrion and J. B. Lasserre. Detecting global optimality and extracting solutions in GloptiPoly. In *Positive polynomials in control*, volume 312, pages 293–310, Springer, Berlin, 2005.
- [10] D. Henrion, J. B. Lasserre, and J. Löfberg. GloptiPoly 3: moments, optimization and semi-definite programming. *Optimization Methods and Software*, 24(4-5):761–779, 2009.
- [11] R. Hettich and K. O. Kortanek. Semi-infinite programming: Theory, methods, and applications. *SIAM Review*, 35(3):380–429, 1993.
- [12] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.
- [13] J. B. Lasserre. *Moments, Positive Polynomials and Their Applications*. Imperial College Press, London, UK, 2009.
- [14] J. B. Lasserre. An algorithm for semi-infinite polynomial optimization. *TOP*, 20(1):119–129, 2012.
- [15] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging Applications of Algebraic Geometry of IMA Volumes in Mathematics and its Applications*, volume 149, pages 157–270. Springer, 2009.
- [16] M. López and G. Still. Semi-infinite programming. *European Journal of Operational Research*, 180(2):491 – 518, 2007.
- [17] Z.-Q. Luo, N. D. Sidiropoulos, P. Tseng, and S. Zhang. Approximation bounds for quadratic optimization with homogeneous quadratic constraints. *SIAM Journal on Optimization*, 18(1):1–28, 2007.
- [18] M. Marshall. Optimization of polynomial functions. *Canadian Mathematical Bulletin*, 46(4):575–587, 2003.
- [19] M. Marshall. *Positive Polynomials and Sums of Squares*. Mathematical Surveys and Monographs, 146. American Mathematical Society, Providence, RI, 2008.
- [20] J. Nie. Discriminants and nonnegative polynomials. *Journal of Symbolic Computation*, 47(2):167–191, 2012.
- [21] J. Nie. Certifying convergence of Lasserre’s hierarchy via flat truncation. *Mathematical Programming, Ser. A*, (142):485–510, 2013.
- [22] J. Nie. An exact Jacobian SDP relaxation for polynomial optimization. *Mathematical Programming, Ser. A*, 137:225–255, 2013.
- [23] P. M. Pardalos and S. A. Vavasis. Quadratic programming with one negative eigenvalue is NP-hard. *Journal of Global Optimization*, 1(1):15–22, 1991.
- [24] P. Parpas and B. Rustem. An algorithm for the global optimization of a class of continuous minimax problems. *Journal of Optimization Theory and Applications*, 141(2):461–473, 2009.
- [25] P. A. Parrilo and B. Sturmfels. Minimizing polynomial functions. In *Algorithmic and quantitative real algebraic geometry*, volume 60 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 83–99. American Mathematical Society, 2003.

- [26] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42(3):969–984, 1993.
- [27] R. Reemtsen and S. Görner. Numerical methods for semi-infinite programming: A survey. In R. Reemtsen and J.-J. Rückmann, editors, *Semi-Infinite Programming*, volume 25 of *Nonconvex Optimization and Its Applications*, pages 195–275. Springer US, 1998.
- [28] W. W. Rogosinski. Moments of non-negative mass. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 245(1240):1–27, 1958.
- [29] K. Schmüdgen. The K-moment problem for compact semi-algebraic sets. *Mathematische Annalen*, 289(1):203–206, 1991.
- [30] A. Shapiro. Semi-infinite programming, duality, discretization and optimality conditions. *Optimization*, 58(2):133–161, 2009.
- [31] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11/12:625–653, 1999.
- [32] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996.
- [33] L. Wang and F. Guo. Semidefinite relaxations for semi-infinite polynomial programming. *Computational Optimization and Applications*, 58(1):133–159, 2013.
- [34] H. Wolkowicz, R. Saigal, and L. Vandenberghe. *Handbook of Semidefinite Programming - Theory, Algorithms, and Applications*. Kluwer Academic Publisher, Dordrecht, 2000.
- [35] Y. Xu, W. Sun, and L. Qi. On solving a class of linear semi-infinite programming by SDP method. *Optimization*, 64(3):603–616, 2015.

SCHOOL OF MATHEMATICAL SCIENCES, DALIAN UNIVERSITY OF TECHNOLOGY, DALIAN, 116024, CHINA

E-mail address: fguo@dlut.edu.cn